1 Order of convergence

Definition 1.1 (From Süli & Mayers). Suppose that
\[ \lim_{k \to \infty} x_k = \xi. \]

If there exists a sequence \( \{\epsilon_k\} \) of positive numbers converging to 0 and \( \mu > 0 \) such that
\[ |x_k - \xi| \leq \epsilon_k, \, k = 0, 1, 2, \ldots, \quad \text{and} \quad \lim_{k \to \infty} \frac{\epsilon_{k+1}}{\epsilon_k} = \mu, \]
then we say that the sequence \( \{x_j\} \) converges to \( \xi \) with at least order \( q > 1 \).

Furthermore, if the above assumptions hold with \( |x_k - \xi| \leq \epsilon_k \) (instead of \( \leq \)), then the sequence is said to converge to \( \xi \) with order \( q \). If \( q = 2 \) then the sequence is said to converge quadratically.

Example 1.1. We proved in class that Newton’s method converges quadratically, and that in fact,
\[ \lim_{k \to \infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|^2} = \frac{1}{2} \left| \frac{f''(\xi)}{f'(\xi)} \right|. \]

Example 1.2. Recall that the secant method for solving \( f(x) = 0 \) is given by:
\[ x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k). \]

The fraction above is a finite difference approximation to the value \( 1/f'(x) \), which is used in Newton’s method. One can show that the secant method converges with order in between 1 and 2, in fact, \( q = (1 + \sqrt{5})/2 \) for the secant method. (This is a homework exercise.)

2 Multi-dimensional Newton’s method

Newton’s method can also be used to solve systems of nonlinear equations. The general two-variable case is given as:
\[ f_1(x_1, x_2) = 0 \]
\[ f_2(x_1, x_2) = 0 \]

This system can be written concisely in vector notation simply as \( \mathbf{f}(\mathbf{x}) = \mathbf{0} \).
To begin, let’s first recall the form of multivariate Taylor series expansions. In the case of a scalar function \( f \) of two variables expanded about the point \( y = (y_1, y_2) \):

\[
f(x) = f(y) + \frac{\partial f}{\partial x_1}(y)(x_1 - y_1) + \frac{\partial f}{\partial x_2}(y)(x_2 - y_2) + \frac{1}{2!} \left( \frac{\partial^2 f}{\partial x_1^2}(y)(x_1 - y_1)^2 + 2\frac{\partial^2 f}{\partial x_1 \partial x_2}(y)(x_1 - y_1)(x_2 - y_2) + \frac{\partial^2 f}{\partial x_2^2}(y)(x_2 - y_2)^2 + \right) + \ldots \quad (1)
\]

For convenience, we will suppress the evaluation point \( y \), it will be implied. The above expression can be written more succinctly as

\[
f(x) = f + (f_{x_1} \ f_{x_2})(x - y) + (x - y)^T \begin{pmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_2x_1} & f_{x_2x_2} \end{pmatrix} (x - y) + \ldots
\]

\[
= f + Df(x - y) + (x - y)^T D^T Df(x - y) + \ldots
\]

The \( 2 \times 2 \) matrix \( D^T Df \) is usually called the Hessian of \( f \).

Now, unfortunately, we need to do a related calculation for a vector-valued function, \( f = f(x) \) instead of a scalar-valued one. In this case, the multivariable Taylor series is just computed for each component of the vector \( f \), which is just a vector of functions:

\[
f(x) = \left( \begin{array}{c} f_1(x) \\ f_2(x) \end{array} \right)
\]

In this case, the first two terms in the vector-valued multivariate Taylor series are:

\[
f(x) = f(y) + J(x - y) + \ldots
\]

where the matrix \( J \) is known as the Jacobian matrix, and is given by:

\[
J = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix}
\]

The higher order terms in the above Taylor expansion formally involve tensors, which will not discuss. With the above expansion, we can derive Newton’s method in the vector-valued case. Recall what Newton’s method in one dimension does: it approximates the nonlinear function \( f \) by a linear one using the tangent line, and then finds the root of the tangent line. This root is used as an approximation to the root, and a new tangent line is formed, and the process repeats.

If we were to truncate the above infinite expansion, we would have:

\[
f(x) \approx f(y) + J(x - y) + \ldots
\]

Furthermore, if we denote the root of \( f \) by \( z \), then we have that

\[
f(z) = 0 \approx f(y) + J(y)(z - y),
\]

2 of 3
where the evaluation point of $J$ is made explicit. Solving the above approximate equation for $z$ we have that:

$$z \approx y - J(y)^{-1}f(y).$$

Therefore, starting with an initial guess for the root $x_0$, we have that the multidimensional Newton’s method is:

$$x_{k+1} = x_k - J(x_k)^{-1}f(x_k).$$

Convergence for the multidimensional Newton’s method is also quadratic (under some assumptions similar to those in the one-dimensional case, which we will not detail). In this case, we must replace the absolute errors with norms:

$$\|x_{k+1} - z\| \approx A\|x_k - z\|$$

where $A$ is some constant, and $\| \cdot \|$ denotes the usual $\ell_2$ Euclidean norm:

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$